

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Administrators of Tulane Educational Fund
Philipp, Mario T.
- (ii) TITLE OF INVENTION: Surface Antigens and Proteins Useful in
Compositions for the Diagnosis and Prevention of Lyme
Disease
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Howson and Howson
(B) STREET: Spring House Corporate Cntr., P.O. Box 457
(C) CITY: Spring House
(D) STATE: Pennsylvania
(E) COUNTRY: USA
(F) ZIP: 19477
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: WO
(B) FILING DATE:
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/051,271
(B) FILING DATE: 30-JUN-1997
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Bak, Mary E.
(B) REGISTRATION NUMBER: 31,215
(C) REFERENCE/DOCKET NUMBER: TUL2APCT
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 215-540-9200
(B) TELEFAX: 215-540-5818

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1047 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA

SECRET "E054460"

T0900

53

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1047

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAG AAT AAT GAT CAT GAT AAT CAT AAG GGG ACT GTT AAG AAT GCT GTT	48
Lys Asn Asn Asp His Asp Asn His Lys Gly Thr Val Lys Asn Ala Val	
1 5 10 15	
GAT ATG GCA AAG GCC GCT GAG GAA GCT GCA AGT GCT GCA AGT GCT GCT	96
Asp Met Ala Lys Ala Ala Glu Glu Ala Ala Ser Ala Ala Ser Ala Ala	
20 25 30	
ACT GGT AAT GCA GCG ATT GGG GAT GTT GTT AAG AAT AGT GGG GCA GCA	144
Thr Gly Asn Ala Ala Ile Gly Asp Val Val Lys Asn Ser Gly Ala Ala	
35 40 45	
GCA AAA GGT GGT GAG GCG GCG AGT GTT AAT GGG ATT GCT AAG GGG ATA	192
Ala Lys Gly Gly Glu Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile	
50 55 60	
AAG GGG ATT GTT GAT GCT GCT GGA AAG GCT GAT GCG AAG GAA GGG AAG	240
Lys Gly Ile Val Asp Ala Ala Gly Lys Ala Asp Ala Lys Glu Gly Lys	
65 70 75 80	
TTG GAT GCT ACT GGT GCT GAG GGT ACG ACT AAC GTG AAT GCT GGG AAG	288
Leu Asp Ala Thr Gly Ala Glu Gly Thr Thr Asn Val Asn Ala Gly Lys	
85 90 95	
TTG TTT GTG AAG AGG GCG GCT GAT GAT GGT GGT GAT GCA GAT GAT GCT	336
Leu Phe Val Lys Arg Ala Ala Asp Asp Gly Gly Asp Ala Asp Ala	
100 105 110	
GGG AAG GCT GCT GCT GCG GTT GCT GCA AGT GCT GCT ACT GGT AAT GCA	384
Gly Lys Ala Ala Ala Ala Val Ala Ala Ser Ala Ala Thr Gly Asn Ala	
115 120 125	
GCG ATT GGA GAT GTT GTT AAT GGT GAT GTG GCA AAA GCA AAA GGT GGT	432
Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly	
130 135 140	
GAT GCG GCG AGT GTT AAT GGG ATT GCT AAG GGT ATA AAG GGG ATT GTT	480
Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val	
145 150 155 160	
GAT GCT GCT GAG AAG GCT GAT GCG AAG GAA GGG AAG TTG AAT GCT GCT	528
Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala	
165 170 175	
GGT GCT GAG GGT ACG ACT AAC GCG GAT GCT GGG AAG TTG TTT GTG AAG	576
Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys	
180 185 190	
AAT GCT GGT AAT GTG GGT GGT GAA GCA GGT GAT GCT GGG AAG GCT GCT	624
Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala	
195 200 205	

666121 E0854460

54

GCT GCG GTT GCT GCT GTT AGT GGG GAG CAG ATA TTA AAA GCG ATT GTT Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val 210 215 220	672
CAT GCT GCT AAG GAT GGT GGT GAG AAG CAG GGT AAG AAG GCT GCG GAT His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp 225 230 235 240	720
CGT ACA AAT CCC ATT GAC GCG GCT ATT GGG GGT GCG GGT GAT AAT GAT Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp 245 250 255	768
GCT GCT GCG GCG TTT GCT ACT ATG AAG AAG GAT GAT CAG ATT GCT GCT Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala 260 265 270	816
GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAA TTT GCT TTG AAG Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys 275 280 285	864
GAT GCT GCT GCT GCT CAT GAA GGG ACT GTT AAG AAT GCT GTT GAT ATA Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile 290 295 300	912
ATA AAG GCT GCT GCG GAA GCT GCA AGT GCT GCA AGT GCT GCT ACT GGT Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly 305 310 315 320	960
AGT GCA GCA ATT GGG GAT GTT GTT AAT GGT AAT GGA GCA ACA GCA AAA Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys 325 330 335	1008
GGT GGT GAT GCG AAG AGT GTT AAT GGC ATT GCT AAG GGA Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly 340 345	1047

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Lys Asn Asn Asp His Asp Asn His Lys Gly Thr Val Lys Asn Ala Val 1 5 10 15
Asp Met Ala Lys Ala Ala Glu Glu Ala Ala Ser Ala Ala Ser Ala Ala 20 25 30
Thr Gly Asn Ala Ala Ile Gly Asp Val Val Lys Asn Ser Gly Ala Ala 35 40 45

66E2E1" E0854460

55

Ala Lys Gly Gly Glu Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile
 50 55 60
 Lys Gly Ile Val Asp Ala Ala Gly Lys Ala Asp Ala Lys Glu Gly Lys
 65 70 75 80
 Leu Asp Ala Thr Gly Ala Glu Gly Thr Thr Asn Val Asn Ala Gly Lys
 85 90 95
 Leu Phe Val Lys Arg Ala Ala Asp Asp Gly Gly Asp Ala Asp Asp Ala
 100 105 110
 Gly Lys Ala Ala Ala Ala Val Ala Ala Ser Ala Ala Thr Gly Asn Ala
 115 120 125
 Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly
 130 135 140
 Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val
 145 150 155 160
 Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala
 165 170 175
 Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys
 180 185 190
 Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala
 195 200 205
 Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val
 210 215 220
 His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp
 225 230 235 240
 Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp
 245 250 255
 Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala
 260 265 270
 Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys
 275 280 285
 Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile
 290 295 300
 Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly
 305 310 315 320
 Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys
 325 330 335
 Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly
 340 345

66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCGCTGGAT GGTGGTGAGA AGCAGGGTAA GAAGGCTGCG GATCGTACAA ATCCCATTGA	60
CGCGGCTATT GGGGGTGCGG GTGATAATGA TGCTGCTGCG GCGTTTGCTA CTATGAAGAA	120
GGATGATCAG ATTGCTGCTG CTATGGTTCT GAGGGGAATG GCTAAGGATG GGCAATTTGC	180
TTTGAAGGAT GCTGCTGCTG CTCATGAAGG GACTGTTAAG AATGCTGTTG ATATAATAAA	240
GGCTGCTGCG GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGT	283

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTATTATAT CAACAGATTC TTAACAGTCC CTTTCATGAGC AGCAGCAGCA TCCTTCAAAG	60
CAAATTGCCC ATCCTTAGCC ATTCCCCTCA GAACCATAGC AGCAGCAATC TGATCATCCT	120
TCTTCATAGT AGCAAACGCC GCAGCAGCAT CATTATCACC CGCACCCCCA ATAGCCGCGT	180
CAATCGGATT TGTACGATCC GCAGCCTTCT TACCCTGCTT CTCACCACCA TCC	233

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

57

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGTGCAAGC TGGGTTGAAG AAGGTTGGGG ATGTTGTAA GAATAGTGAG GCAAAGATG	60
GTGATGCGGC GAGTGTAAAT GGGATTGCTA AGGGGATAAA GGGGATTGTT GATGCTGCTG	120
AGAAGGCTGA TCGAAGGAA GGAAGTTGG TATGTGGCTG GTGCTGCTGG TGAACTAAC	180
AAGGAAGCGG CCGC	194

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGGCCGCTT GAGGAAGCTG CAAGTGCTGC AAGTGCTGCT ACTGGTAATG CAGCGATTGG	60
GGATGTTGTT AAGAATAGTG GGCAGCAGC AAAAGGTGGT GAGGCGGCGA GTGTTAATGG	120
GATTGCTAAG GGGATAAAGG GGATTGTTGA TGCTGCTGGA AAGGCTGATG CGAAGGAAGG	180
GAAGTTGGAT GCTACTGGTG CTGAGGGTAC GACTAACGTG AATGCTGGGA AGTTGTTTGT	240
GAAGAGGGCG GCTGATGATG GTGGTGATGC AGATGATGCT GGAAGGCTG CTGCTGCGGT	300
TGCTGCAAGT GCTGCTACTG GTAATGCAGC GATTGGAGAT GTTGTTAATG GTGATGTGGC	360
AAAACAAA	369

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 142 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGGATGGTG ATGATAAGCA GGGTAAGAAG GCTGAGGATG CTACAAATCC GATTGACGCG	60
GCTATTGGGG GTGCAGGTGC GGGTGCTAAT GCTGCTGCGG CGTTTAATAA TATGAAGAAG	120
GATGATCAGA TTGAGCGGCC GC	142

09445803 "121393
 66EET" 6854460

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGTGAACTA ACAAGGATGC TGGGAAGTTG TTTGTGAAGA AGAATGGTGA TGATGGTGGT 60
GATGCAGGTG ATGCTGGGAA GGCTGCTGCT GCGGTTGCTG CTGTTAGTGG GGAGCAGATA 120
TTAAAAGCGA TTGTTGATGC TGCTAAAGAT GGTGATAAGA CGGGGGTTAC TGATGTAAAG 180
GATGCTACAA ATCCGATTGA CGCGGCTATT 210

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 236 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATATAATAA AGGCTGCTGC GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGTAGTGCAG 60
CAATTGGGGA TGTTGTTAAT GGTAATGGAG CAACAGCAAA AGGTGGTGAT GCGAAGTGTT 120
AATGGGATTG CTAAGGGGAT AAAGGGGATT GTTGATGCTG CTGAGAAGGC TGATGCGAAG 180
GAAGGGAAGT TGGATGTGGC TGGTGATGCT GGTGAACTA ACAAGGAAGC GGCCGC 236

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGAGAGGAT CTCATCACCA TCACCATCAC ACGGATCCCC CGGGCTGCAG GAATTCGCGG 60
CCGCTGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGCTGC TGGTGAACT 120

AACAAGGATG CTGGGAAGTT GTTTGTGAAG AAGAATAATG AGGGTGGTGA AGCAAATGAT 180
GCTGGGAAGG CTGCTGCTG 199

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 272 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCCTGGAT GATCAGATTG CTGCTGCTAT GGTGTGAGG GGAATGGCTA AGGATGGGCA 60
GTTTGCTTTG AAGGATGATG CTGCTAAGGA TGGAGATAAA ACGGGGGTTG CTGCGGATGT 120
GAAAATCCGA TTGACGCGGC TATTGGGGGT GCGGATGCTG ATGCTGCGGC GTTTAATAAG 180
GAGGGGATGA AGAAGGATGA TCAGATTGCT GCTGCTATGG TTCTGAGGGG AATGGCTAAG 240
GATGGGCAGT TTGCTTTGAC GAATAATGCT GC 272

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 289 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTGTTAAGA ATGCTGTTGA TATAATAAAG GCTGCTGCGG AAGCTGCAAG TGCTGCAAGT 60
GCTGCTACTG GTAGTGACGC AATTGGGGAT GTTGTTAATG GTAATGGAGC AACAGCAAAA 120
GGTGGTGATG CGAAGAGTGT TAATGGGATT GCTAAGGGGA TAAAGGGGAT TGTTGATGCT 180
GCTGAGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGATGC TGGTGAAACT 240
AACAAGGATG CTGGGAAGTT GTTTGTGAAG AACAATGGTA ATGAGGGTA 289

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

60

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

G CCG CTT ACA AAT CCG ATT GAC GCG GCT ATT GGG GGG AGT GCG GAT	46
Pro Leu Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ser Ala Asp	
1 5 10 15	
CGT AAT GCT GAG GCG TTT GAT AAG ATG AAG AAG GAT GAT CAG ATT GCT	94
Arg Asn Ala Glu Ala Phe Asp Lys Met Lys Lys Asp Asp Gln Ile Ala	
20 25 30	
GCT GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAG TTT GCT TTG	142
Ala Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu	
35 40 45	

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro Leu Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ser Ala Asp Arg	
1 5 10 15	
Asn Ala Glu Ala Phe Asp Lys Met Lys Lys Asp Asp Gln Ile Ala Ala	
20 25 30	
Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu	
35 40 45	

666T2T 605460